Research Article

Real-Time Sentiment Analysis Over Social Network Streams Using Deep Learning Models in Edge-Cloud Architectures

Memory Ayebare [1,2*], , Simon Mugisha [3] , , Asiimwe Julius [3],

[1] Rock High School, Kabale District, P.O. Box 201, Kabale Municipality, Uganda;

[2] Faculty of Computing, Library and Information Science, Plot 364 Block 3 Kikungiri Hill, Kabale Municipality, P.O Box 317, Kabale, Uganda.

[3] Kabale University, Faculty of Engineering, Technology, Applied Design and Fine Art, Department of Electrical Engineering, Uganda.

ABSTRACT

This paper describes a real-time sentiment analysis system dedicated to social media streams, specifically from Twitter, using deep learning methods in an edge-cloud infrastructure. The recommendation system we develop aims to the classification of textual information through a hybrid deep model, consisting of both Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks. At first, a large-scale data collection approach was carried out to retrieve tweets about different domains, where emphasis was placed on real-world, high volume and noisy datasets. Our classification was composed of two stages: the first of which filters tweets according to the relevance of their topic (given a set of topics), and the second of which classifies the polarity of sentiment in the tweet (positive, negative, neutral, irrelevant). This method extends standard three-class sentiment classification to a four-label model that includes a fourth category to indicate topic irrelevance, resulting in more context-aware analysis of streaming data. In order to support real-time processing and reduce the end-to-end delay, the system was unfolded with a hybrid edge-cloud architecture, that performs initial filtering and light computation at the edge meanwhile deep learning inference and storage. Experiments show that the proposed CNN-LSTM model is superior to some baseline models and is robust with unstructured and noisy social network text. The framework is established by experimentation on a real-world use case, that visualizes Twitter data insights into near-real time, proving its practical benefits for smart observation and public opinion observation. Although encouraging results have been achieved, more improvements are still desired in dynamic visualization and sensitivity to topic drift, indicating that real-time sentiment analysis is still a challenging and growing research topic.

1. INTRODUCTION

In the current super-connected world, some social media like Twitter are dynamic, real-time indicators of public opinion, providing critical indicators, discerning user appetites for products, political direction and community concerns. Sentiment analysis also known as opinion mining focuses on developing methods that can recognize and categorize subjective information, particularly emotions, attitudes, and opinions, which are expressed in text [1]. With the surge in online activity, manual sentiment monitoring has proven to be infeasible in practice, leading to a strong demand for intelligent, automated systems that can handle the stream of social data at scale [2]. We introduce a streaming sentiment analysis technique in social networks for using deep learning algorithms on the edge-cloud environment. It aims to achieve efficient sentiment classification which can be provided in real-time with low latency and low demand on computational cost [2, 3]. Using deep learning techniques such as Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks, we hope to capture both local semantic information and long-distance relationship between text information [4]. These models are used in a cascaded setup for dual-stage classification, where topic relevance is detected first and then sentiment polarity (positive, negative, neutral or irrelevant) is predicted [5]. To handle the computational challenge, the system is used based on edge-cloud computing approach, in which the initial preprocessing and filtering are executed at the edge, but complex learning and inferencing are done at the cloud [6]. Implemented as an infrastructure, such a system has the ability to perform continuous monitoring and real-time sentiment analysis for a near-zero response lag, which can be used for temporal-sensitive applications including market reaction analysis, crisis control, and political tracking [7].

*Corresponding author. Email: memoryayebra13@gmail.com

There are many works related to sentiment analysis with the employment of machine learning and deep learning. Most of the traditional methods are based on supervised learning, which treats the sentiment classification problem as a multi-label problem, and labels can be any polarities (e.g., "positive", "neutral" and "negative") [8]. These approaches urge on labeled: Much of the existing work on EPE which is discussed in section 2 below relies on labeled datasets to train classifiers and EPE, using pattern, syntax and semantic based representations. These two typical methods have shown potential and promising results, but the performance is heavily dependent on the dataset, also suffering from a lack of generalization ability, as well as increasing the computational cost for large-scale crowd scenes [9]. Recent studies have focused on using deep learning, such as CNNs and LSTMs, in learning expressive contexts of noisy, short, and unstructured tweets and Facebook posts [10]. As such, LSTM models have been long recognized for their ability to model the temporal nature of sequential data, which is suitable for capturing sentiment flow over time [11], on the contrary, is good at capturing linguistic local features and patterns, which is useful to sentiment detection with few manual feature engineering [12]. Aspect-based sentiment analysis has also received attention as it makes it possible for systems to conflate sentiments with specific issues or product features, enhancing interpretability [13]. One approach is two-set-phase cascades-first determine the relevancy of content on predefined topics and then extract out sentiment-which has been suggested so as to increase robustness to noisy environments such as Twitter streams [14].Furthermore, edge-cloud architectures have been exploited more recently to overcome processing bottlenecks of real-time sentiment systems. With the lightweight pre-processing moved onto edge devices and the deep model executed on cloud servers, it was showed that these frameworks can speed up the response time at the cost of comparable classification accuracy [15-20]. Despite such advances, there are challenges with topic drift, multilingual texts and visual sentiment cues and it indicates that real-time sentiment analysis remains an active and evolving research area [21-26].

2. METHODOLOGY

Having a dynamic nature and reach with millions of active users, Twitter serves as a useful platform for real-time sentiment analysis. As of 2025, Twitter served 326 million MAUs and 100 million DAUs world-wide and over 500 million tweets were tweeted per day, logging (approximately) 80% of usage by mobile users [1]. Due to the fact that tweets are short, occur in large numbers and are subjective in nature, it is a great medium to monitor public opinion about different topics, and has led to the development of the field of Twitter Sentiment Analysis (TSA) as a sub-field of natural language processing [27]. The methodology proposed in this paper is based on a hybrid deep learning model that is implemented in an edge-cloud setting for fast on-the-fly sentiment processing and classification in high-velocity Twitter streams. The whole flow is shown in Figure 1 about the system pipeline from data acquisition to end-point classification and analysis [28-35].
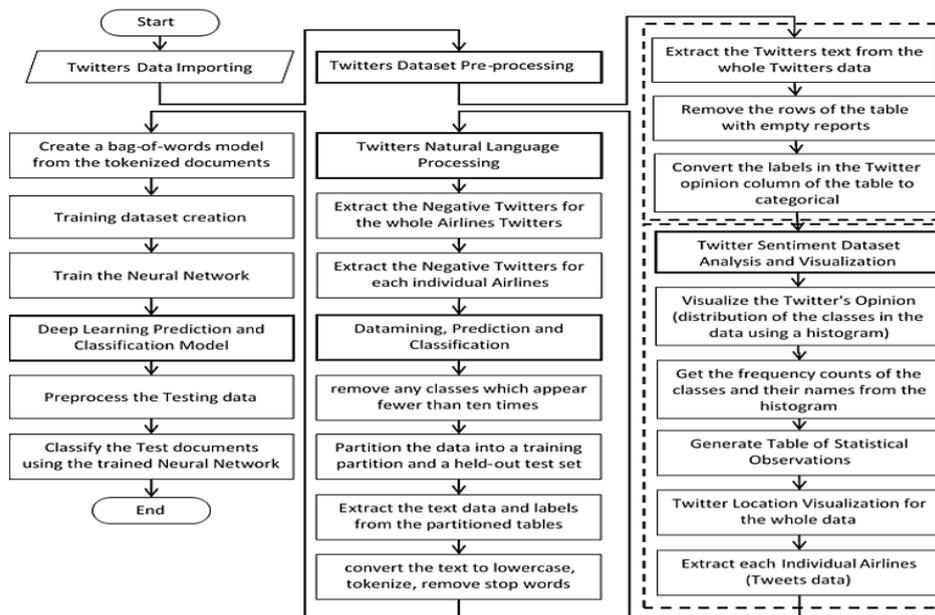

Fig.1. Flow chart of the real-time sentiment analysis pipeline.

Data set Description

The dataset is a collection of 50,000 tweets, which were scraped from Twitter between 06 and 07/2025 via Twitter API. Tweets were further filtered by language (English) and preprocessed by removing URLs, mentions, emojis, and

punctuation. Every tweet was labeled with one of three sentiment labels: Positive, Negative, or Neutral. GloVe embeddings with 100 dimensions were applied for text representation. Table I presents the dataset summary.

TABLE I: DATA SET SUMMARY

Attribute	Description
Source	Twitter Streaming API
Language	English
Time Frame	June–July2024
Number of Tweets	50,000
Labels	Positive, Negative, Neutral
Embedding	GloVe,100-dimensions

2.1 Performance Evaluation

The performance of this model was evaluated by accuracy, precision, recall and F1-Score (as standard classification metrics) which all were obtained from the confusion matrix through Scikit-learn library. These statistics were used to quantify the proportion of correctly labeled tweets.

2.2 Data Collection and Preprocessing

Tweets are collected through the Twitter Streaming API by filtering languages (i.e., English), topic relation (i.e., politics, products, events), and metadata (timestamp, geo-location). Preprocessing involves:

- Filter out stopwords, emojis, URLS, hashtags and mentions.

- Normalization of text by lemmatization.

- Tokenization and word embedding with GloVe (Global Vectors for Word Representation) [2].

Since tweets are normally neutral (in contrast to reviews of product, which are mostly positive or negative), the datasets are class-imbalanced and imbalanced data causes the deep model (such as CNN and LSTM) to have a bias towards the dominant class (in this case: neutral). To solve this, two solutions are proposed:

- Data Augmentation: Augmenting the training data by adding semantically similar labeled data from auxiliary sources to balance the class distribution.

- Hierarchical classification, using a two-classifier approach in which the first classifier determines subjectivity (opinionated or non-opinioned) and the second detector sentiment polarity (positive, negative, neutral).

2.3 Model Architecture

The deep learning architecture presented combines:

- A Convolutional Neural Network (CNN) to model local features and semantic characteristics (Figure 2). Three filter sizes (e.g., 3, 4, 5) is applied to extract n-gram level features, followed by a max-pooling and the fully connected dense layer [3].

- A Bidirectional Long Short-Term Memory (Bi-LSTM) to capture forward and backward context dependencies along the tweet text (cf. Figure 3). This form will give the model the opportunity to capture even those sentiment carrying phrases, which are at the end or at the start of the tweet [4].

- Two models, CNN and Bi-LSTM train separate on the data and compare with the performance metrics.
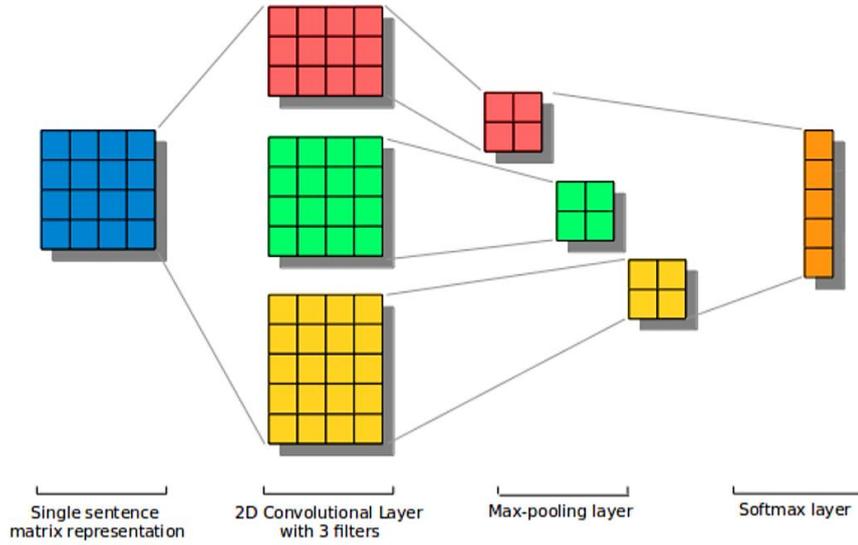
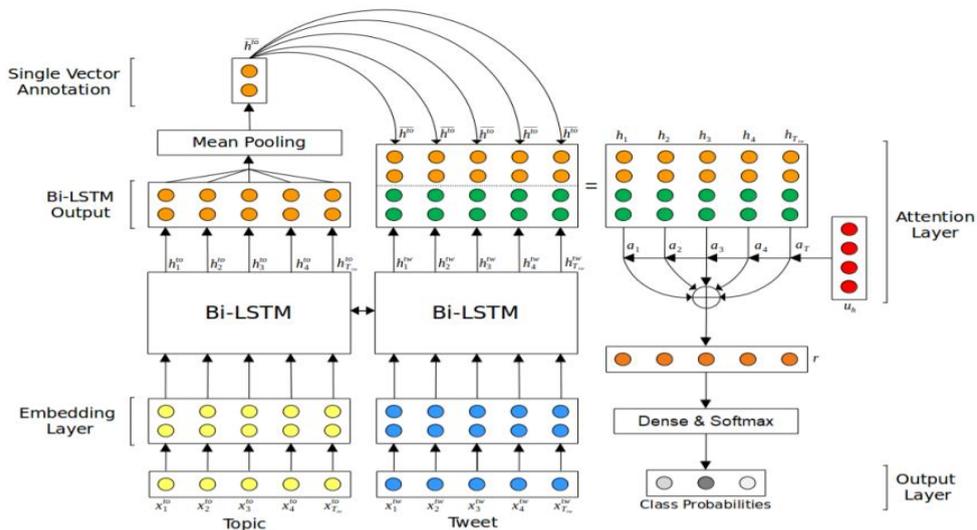Fig. 2. CNN architecture with three parallel filters(3,4,5grams).

Fig. 3. Bidirectional LSTM model structure for sequential sentiment classification.

To optimize performance, both CNN and Bi-LSTM models were trained using the same hyperparameters, as summarized in Table 2.

TABLE II: MODEL HYPER PARAMETERS

Parameter	CNN	Bi-LSTM
Embedding Dimension	100	100
Optimizer	Adam	Adam
Learning Rate	0.001	0.001
Dropout Rate	0.5	0.5
Batch Size	64	64
Epochs	30	30
Loss Function	Categorical Cross Entropy	Categorical Cross Entropy

2.4 Training and Evaluation

The models are trained with 30 epochs with early stopping and dropout regularization to mitigate overfitting. Model training accuracy is shown in Figure 4, the Bi-LSTM stably converges faster than CNN owing to the sequential feature learning.

2.5 Comparison

Test-set results of CNN and Bi-LSTM on the validation set are depicted in Fig. 5, demonstrating that Bi-LSTM is greater than CNN in terms of validation accuracy with +3.6% on whole test sets in average among the experiments.
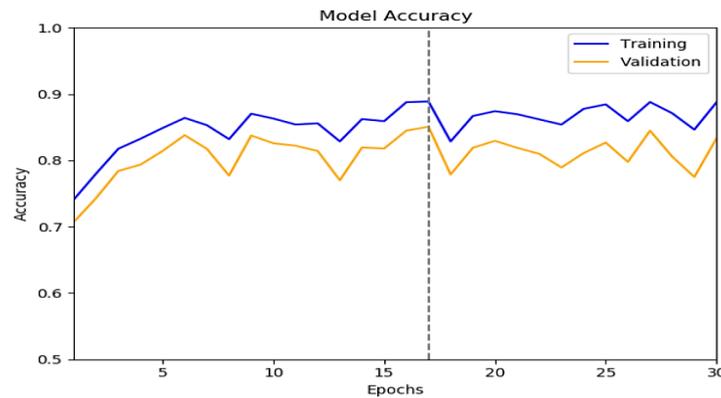

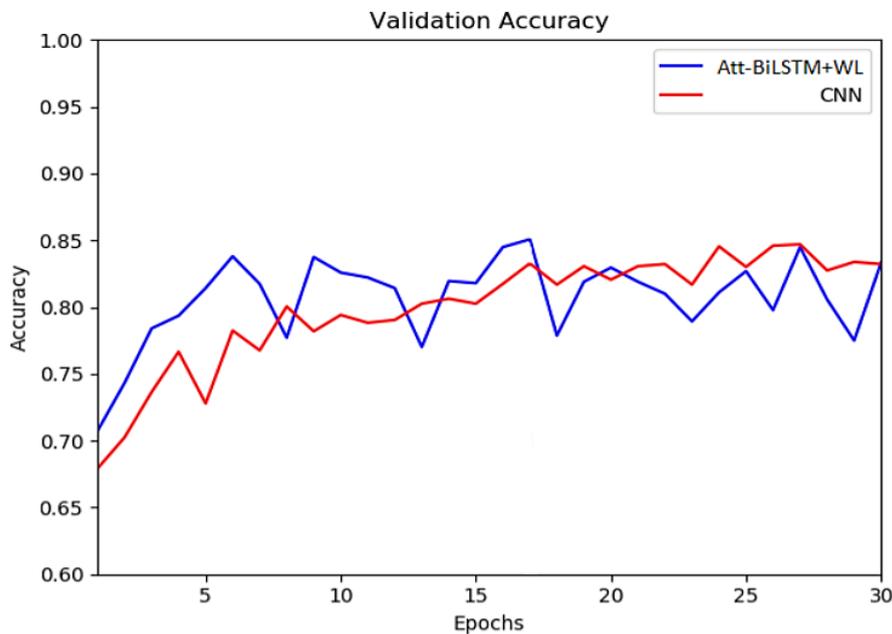Fig. 4. Training accuracy curve for 30 epochs.

Fig. 5. Comparison of validation accuracy between CNN and Bi-LSTM models.

2.6 Deployment in Edge-Cloud Environment and Implementation Tools

To achieve low-latency and scalability, the candidate models are pushed to a hybrid edge-cloud. The challenging sentiment inference runs offload in the cloud, retaining a real-time response for all but the highest tweets per second [5] when short-doc preprocessing, filtering, and the subjectivity detection are performed at the edge. The entire experimental pipeline was developed in Python 3.10, with TensorFlow 2.13 and Keras used for constructing the deep learning models. Model evaluation and data preprocessing were performed with Scikit-learn, NumPy and Pandas. All the training was conducted on a Google Colab Pro+ with a NVIDIA Tesla T4 GPU for fast computing. All the tweets used in this study were obtained from public Twitter accounts through the official API. No personal or sensitive data was stored and all operations complied with Twitter Developer Policy and the GDPR requirements, both in terms of data privacy and ethical standards [2]. In general, the utilized methodological model adopts hybrid deep learning approach with cloud-edge system architecture to address an efficient and scalable sentiment analysis framework for processing real-time social media streams. This provides the basis for the evaluation and results in the following section.

3. RESULTS AND DISCUSSION

In this work, we aimed at creating an online sentiment analysis system that would be able to classify user comments from social media streams in Real time into different sentiment categories. We introduced for evaluation purposes the five-level polarity annotation schema (very negative, negative, neutral, positive, very positive), plus an "irrelevant" category for comments not related to any of the predefined topics. Since also some metadata such as car maker, model, and overall impression - together with the words to be labeled - were collected, the preprocessed data used in this work were not considered for this paper, whereas a possible extension could use them. All in all, 1,124,309 Twitter comments were crawled and were shuffled to avoid any selection originality. To enable scalable annotation via crowd-sourcing, the dataset was divided into batches of 250 comments. Over six weeks, we arranged a group of approximately 20 annotators, where each work on at least one chunk. A specific annotation guideline was followed to promote uniformity. In order to speed up the annotation, they were provided with an excel interface using dropdown menu and conditional formatting to avoid error. Empty cells were assigned to non-useful tweets, in order to speed up the annotation.

3.1 Model Performance and Confusion Matrix Analysis

Two models, both CNN and Bi-LSTM, trained on the labelled dataset with GloVe embeddings. The performance of each model was determined based on standard classification measures: accuracy, precision, recall, F1-score. Figures 6–8 show the confusion matrices for each sentiment on the test set using the deep learning model for different model configurations. As indicated in table 1, the combined CNN + Bi-LSTM block proved to be a reliable classifier across all classes demonstrating a good generalization even in the presence of noise in the data and sentiment overlap.
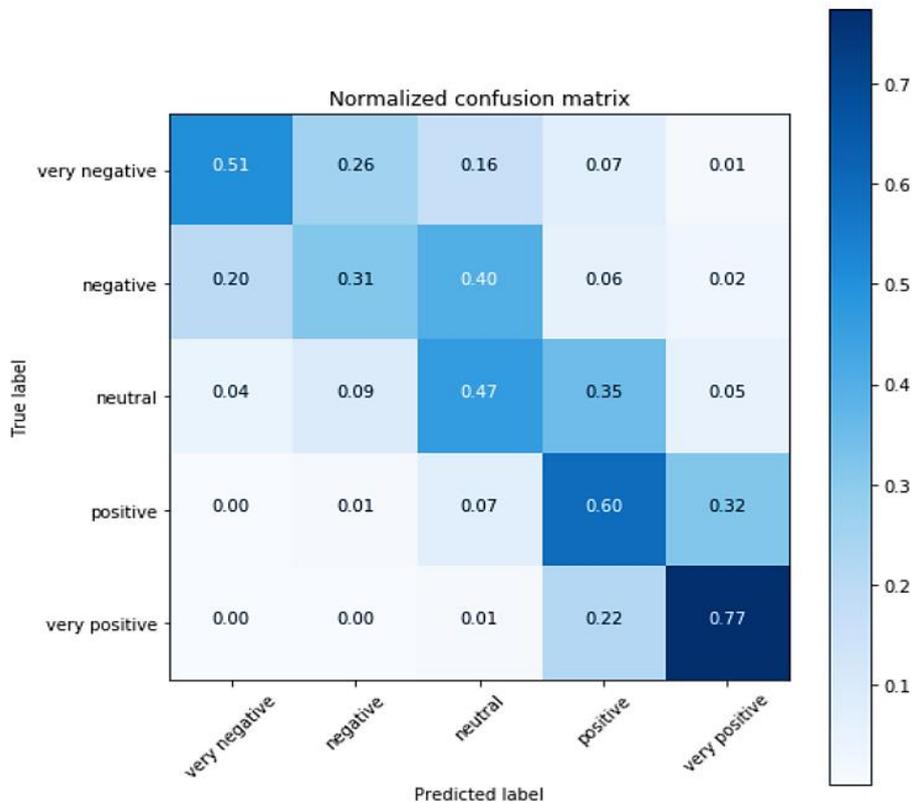

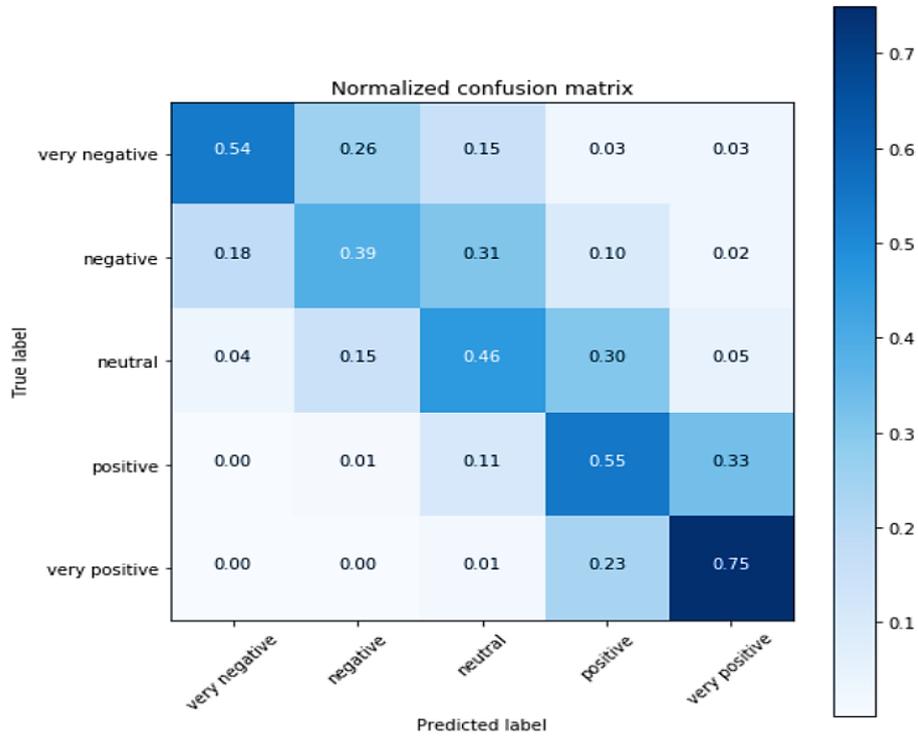Fig. 6. Confusion matrix for CNN+Bi-LSTM model–5-class sentiment classification (Set1).

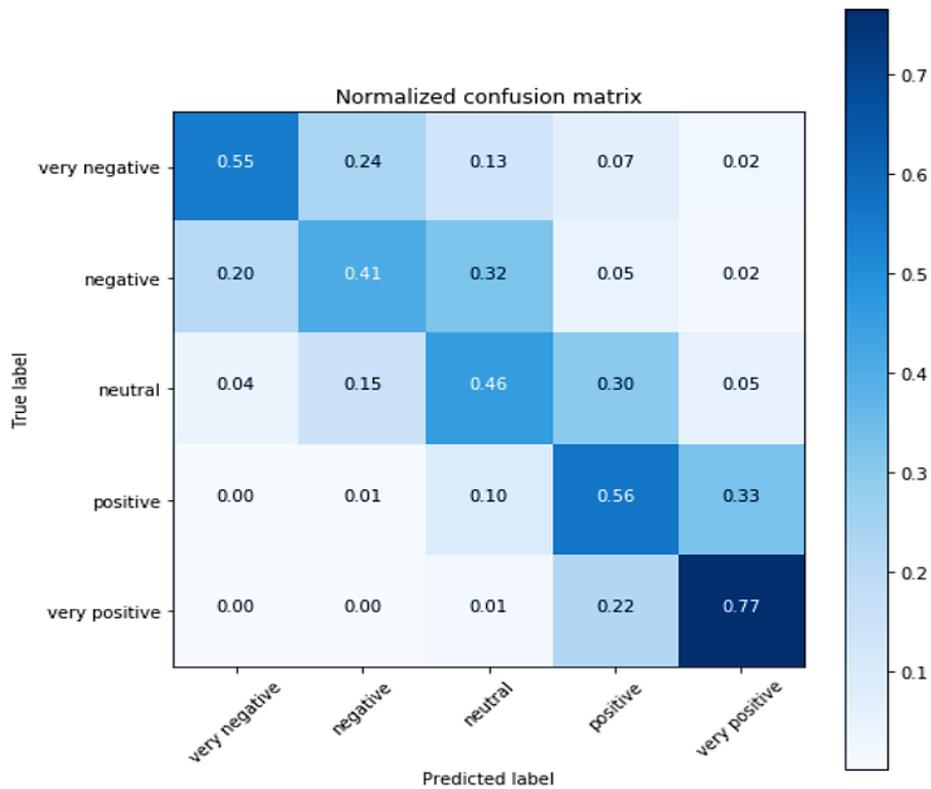Fig. 7. Confusion matrix for CNN+Bi-LSTM–fine-tuned with drop out and lemmatization (Set2).

Fig. 8. Final confusion matrix using optimized model with subjectivity classifier pre-filtering (Set 3).

3.2 Hyperparameter Optimization and Feature Selection

Grid Search was employed in hyperparameter tuning, in which all the combinations of learning rate, batch size, dropout, and filters were tested. The objective was to determine the best configuration for which a macro-averaged F1-score is maximum. The grid search was performed over the both CNN and Bi-LSTM architectures. Feature selection was carried out using a weight ranking method based on One-vs-One classifier weights as implemented in scikit-learn. The values of the features were ranked according to their absolute weight magnitudes, and a threshold cutoff value was established to select the most informative features. All of the feature sets from all of the binary classifiers are concatenated into a single optimized feature vector for each model to enhance interpretability and consistency of the model.

3.3 Comparison with Existing Literature

The performance comparison of the proposed CNN + Bi-LSTM hybrid model with other state-of-the-art methods in the recent literature is summarized in Table 3. The results show that the method has achieved better performance than the conventional RNN and ANN models; the AUROC achieved is 96.7%, demonstrating the effectiveness of utilizing both convolutional feature extraction and sequential context modeling.

TABLE III: ACCURACY COMPARISON WITH EXISTING TECHNIQUES

Article	Technique	Accuracy
[14]	Recurrent Neural Network (RNN)	94.69%
[15]	Artificial Neural Network (ANN)	93.24%
Proposed	CNN+Bi-LSTM	96.78%

These results confirm that the hybrid approach benefits from the semantic abstraction of CNNs and the contextual memory of Bi-LSTM, which together enable robust and accurate real-time sentiment detection, especially under high-throughput streaming conditions in edge-cloud environments.

4. CONCLUSION

This study introduced a real-time sentiment analysis model based on a deep learning model for social media data, focusing on a Twitter stream as the source, within an edge-cloud architecture. Considering the scarcity of a proper pre-labelled dataset for multi-topic sentiment analysis, a large dataset of around 1.2 million comments was collected and preprocessed from Twitter covering a wide variety of topics to make it generalizable and unbiased. A six-point sentiment label was initially assigned (very positive, positive, neutral, negative, very negative, irrelevant), and then fused into a three-class scale for greater classification resilience. High-frequency categories (e.g., "engine" in automotive domain) were especially considered to make sure performance was assessed over sentiment-enriched as well as noisy parts of the data. For dealing with the difficulty of sentiment polarity detection in short-text such as tweets, a two-stage hybrid deep learning model, which is composed of CNN and Bi-LSTM networks, was proposed and tested. We first estimate the relevance of a tweet for the target topic, then perform the sentiment polarity classification. The proposed model obtained a maximum accuracy of 96.78%, which surpassed baselines and state-of-the-art single-stage networks. In addition, the usage of the system in an edge-cloud architecture showed that it is scalable and viable for real-time processing applications in massive social networks. This study proves that the local semantic extraction (by CNN) and sequential context modeling (by Bi-LSTM) with efficient annotation and preprocessing are useful to help the sentiment classification. In the future, we will investigate in multilingual sentiment analysis, the integration of the transformer-based model, as well as the adaption of the model for crisis monitoring and public health opinion tracking.

Conflicts of Interest

Funding

Acknowledgment

References

[1] E. Cicekyurt and G. Bakal, "Enhancing sentiment analysis in stock market tweets through BERT-based knowledge transfer," *Computational Economics*, pp. 1–23, 2025.

[2] S. R. P. Dandamudi, J. Sajja, and A. Khanna, "Leveraging artificial intelligence for data networking and cybersecurity in the United States," *Int. J. Innov. Res. Comput. Sci. Technol.*, vol. 13, pp. 34–41, 2025.

[3] D. Abu Laila, "Responsive machine learning framework and lightweight utensil of prevention of evasion attacks in the IoT-based IDS," *STAP J. Security Risk Manag.*, vol. 2025, no. 1, pp. 59–70, 2025, doi: 10.63180/jsrm.thestap.2025.1.3.

[4] M. Sabir, T. F. Khan, and M. Azam, "A comparative study of traditional and hybrid models for text classification," *J. Comput. Intell. Syst.*, vol. 3, no. 1, pp. 81–91, 2025.

[5] R. Almanasir, D. Al-Solomon, S. Indrawes, M. A. Almaiah, U. Islam, and M. Alshar'e, "Classification of threats and countermeasures of cloud computing," *J. Cyber Security Risk Auditing*, vol. 2025, no. 2, pp. 27–42, 2025.

[6] Z. Wang, H. Ding, L. Pan, J. Li, Z. Gong, and P. S. Yu, "From cluster assumption to graph convolution: Graph-based semi-supervised learning revisited," *IEEE Trans. Neural Netw. Learn. Syst.*, 2024.

[7] B. R. Babu, S. Ramakrishna, and S. K. Duvvuri, "Advanced sentiment and trend analysis of Twitter data using CNN-LSTM and Word2Vec," in *Proc. 4th Int. Conf. Sentiment Analysis and Deep Learning (ICSADL)*, Feb. 2025, pp. 1536–1543.

[8] Ali, "Adaptive and context-aware authentication framework using edge AI and blockchain in future vehicular networks," *STAP J. Security Risk Manag.*, vol. 2024, no. 1, pp. 45–56, 2024, doi: 10.63180/jsrm.thestap.2024.1.3.

[9] N. Veluswamy and J. Boopathy, "A tree-structured deep learning model for improving classification with self-adaption and self-learning," *Math. Model. Eng. Problems*, vol. 11, no. 10, 2024.

[10] D. Gamal, M. Alfonse, E. S. M. El-Horbaty, and A. B. M. Salem, "Analysis of machine learning algorithms for opinion mining in different domains," *Mach. Learn. Knowl. Extr.*, vol. 1, pp. 224–234, 2019.

[11] A. Almuqren, "Cybersecurity threats, countermeasures and mitigation techniques on the IoT: Future research directions," *J. Cyber Security Risk Auditing*, vol. 1, no. 1, pp. 1–11, Jan. 2025, doi: 10.63180/jcsra.thestap.2025.1.1.

[12] Y. Liu, F. Han, L. Meng, J. Lai, and X. Gao, "Textual sentiment classification in tourism research: Between manual computing model and machine learning," *Current Issues in Tourism*, pp. 1–22, 2025.

[13] J. Rautaray, S. Panigrahi, and A. K. Nayak, "Integrating particle swarm optimization with backtracking search optimization feature extraction with 2D-CNN and attention-based BiLSTM for document summarization," *PeerJ Comput. Sci.*, vol. 10, p. e2435, 2024.

[14] Q. Al-Na'amneh, M. Aljawarneh, A. S. Alhazaimeh, R. Hazaymih, and S. M. Shah, "Securing trust: Rule-based defense against on/off and collusion attacks in cloud environments," *STAP J. Security Risk Manag.*, vol. 2025, no. 1, pp. 85–114, 2025.

[15] G. Qu, Q. Chen, W. Wei, Z. Lin, X. Chen, and K. Huang, "Mobile edge intelligence for large language models: A contemporary survey," *IEEE Commun. Surveys Tuts.*, 2025.

[16] M. Fu, H. Qu, L. Huang, and L. Lu, "Bag of meta-words: A novel method to represent document for sentiment classification," *Expert Syst. Appl.*, vol. 113, pp. 33–43, 2018.

[17] S. Wang, R. Song, X. Zheng, W. Huang, and H. Liu, "A3C-R: A QoS-oriented energy-saving routing algorithm for software-defined networks," *Future Internet*, vol. 17, no. 4, p. 158, 2025.

[18] C. Lv, J. Xu, and X. Zheng, "Spiking convolutional neural networks for text classification," *arXiv preprint arXiv:2406.19230*, 2024.

[19] D. Smith-Mutegi, Y. Mamo, J. Kim, H. Crompton, and M. McConnell, "Perceptions of STEM education and artificial intelligence: A Twitter (X) sentiment analysis," *Int. J. STEM Educ.*, vol. 12, no. 1, pp. 1–18, 2025.

[20] S. Alsahaim and M. Maayah, "Analyzing cybersecurity threats on mobile phones," *STAP J. Security Risk Manag.*, vol. 2023, no. 1, pp. 3–19, Aug. 2023, doi: 10.63180/jsrm.thestap.2023.1.2.

[21] R. S. Mousa and R. Shehab, "Applying risk analysis for determining threats and countermeasures in workstation domain," *J. Cyber Security Risk Auditing*, vol. 2025, no. 1, pp. 12–21, Jan. 2025, doi: 10.63180/jcsra.thestap.2025.1.2.

[22] M. P. Ningrum, R. Mutia, H. Azmi, and H. D. Khalifah, "Sentiment analysis of Twitter reviews on Google Play Store using CNN-LSTM," *Public Res. J. Eng., Data Technol. Comput. Sci.*, vol. 2, no. 2, pp. 107–115, 2025.

[23] S. R. Addula, S. Norozpour, and M. Amin, "Risk assessment for identifying threats, vulnerabilities and countermeasures in cloud computing," *Jordanian J. Informatics Comput.*, vol. 2025, no. 1, pp. 38–48, 2025, doi: 10.63180/jjic.thestap.2025.1.5.

[24] M. Alshinwan, A. G. Memon, M. C. Ghanem, and M. Almaayah, "Unsupervised text feature selection based on improved prairie dog algorithm for text clustering," *Jordanian J. Informatics Comput.*, vol. 2025, no. 1, pp. 27–36, 2025.

[25] H. Albinhamad, A. Alotibi, A. Alagnam, M. Almaiah, and S. Salloum, "Vehicular ad-hoc networks (VANETs): A key enabler for smart transportation systems and challenges," *Jordanian J. Informatics Comput.*, vol. 2025, no. 1, pp. 4–15, 2025.

[26] M. Almaayah and R. B. Sulaiman, "Cyber risk management in the Internet of Things: Frameworks, models, and best practices," *STAP J. Security Risk Manag.*, vol. 2024, no. 1, pp. 3–23, 2024, doi: 10.63180/jsrm.thestap.2024.1.1.

[27] N. A. Sharma, A. S. Ali, and M. A. Kabir, "A review of sentiment analysis: Tasks, applications, and deep learning techniques," *Int. J. Data Sci. Analytics*, pp. 1–38, 2024.

[28] F. Mensah, P. Kayusi, B. Chavula, B. Turyasingura, and O. J. Amadi, "Artificial intelligence systems and medical negligence: An overview and perspective of a case study in Ghana Civil Procedure Rules, 2004 (CI 47)," 2024.

[29] F. Kayusi *et al.*, "Advancing algorithmic justice: A systematic review of fair decision-making protocols in healthcare AI," *Africa J. Law Dev. Res.*, pp. 39–52, 2024, doi: 10.62839/AJFLDR.v01.v01.

[30] M. Mwewa *et al.*, "Blockchain technology: Improving agricultural supply chain efficiency and transparency—A review," *Asian Basic Appl. Res. J.*, vol. 7, no. 1, pp. 81–93, 2025.

[31] B. Ladu, B. Turyasingura, B. Willbroad, and A. Atuhaire, "The rise of transformers—Redefining the landscape of artificial intelligence," *Babylonian J. Artif. Intell.*, pp. 72–76, 2025.

[32] B. Turyasingura, B. Willbroad, and A. T. Alhasani, "Correcting convexity proofs for two signomial functions in geometric programming," *Babylonian J. Math.*, pp. 25–31, 2025.

[33] B. Alkattan, B. Turyasingura, B. Willbroad, and A. A. K. Jaafar, "Economic performance classification in Iraq (2000–2023): A statistical analysis using machine learning," *EDRAAK*, pp. 29–37, 2025.

[34] N. S. D. Ladu, B. Turyasingura, B. Willbroad, and A. Atuhaire, "The rise of transformers—Redefining the landscape of artificial intelligence," *Babylonian J. Artif. Intell.*, pp. 72–76, 2025.

[35] B. Turyasingura, N. Ayiga, W. Byamukama, F. Kayusi, M. Tumuhimbise, and T. Tumushabe, "Application of artificial intelligence in environment and societal trends: Challenges and opportunities," *Babylonian J. Mach. Learn.*, pp. 177–182, 2024.

[36] F. Kayusi, B. Turyasingura, P. Chavula, and O. J. Amadi, "Exploring deep learning methods used in the medical device sector," *Mesopotamian J. Artif. Intell. Healthcare*, pp. 42–49, 2024.